

tubes

Kelompok 1 RA

2025

Penerapan Rantai Markov untuk Menentukan Probabilitas Kategori Hujan Harian di ITERA

```
library(readr) # Untuk membaca CSV
library(dplyr) # Untuk manipulasi data (filter)

##

## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##   filter, lag

## The following objects are masked from 'package:base':
##   intersect, setdiff, setequal, union

library(lubridate) # Untuk bekerja dengan tanggal (memfilter tahun)

## Warning: package 'lubridate' was built under R version 4.2.3

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##   date, intersect, setdiff, union

library(markovchain) # Untuk analisis Rantai Markov

## Warning: package 'markovchain' was built under R version 4.2.3

## Package: markovchain
## Version: 0.9.5
## Date: 2023-09-24 09:20:02 UTC
## BugReport: https://github.com/spedygiorgio/markovchain/issues

##
## Attaching package: 'markovchain'

## The following object is masked from 'package:lubridate':
##   period

library(knitr)
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.2.3

library(forcats)

data <- "D:/College/9th/SemStok/tubes/data_harian_terklasifikasi.csv"

# Membaca data lengkap
data_full <- read_csv(data, show_col_types = FALSE)

# Filter Tahun 2024
data_2024 <- data_full %>%
  mutate(DateTime = ymd(DateTime)) %>% # Konversi kolom ke format Tanggal
  filter(year(DateTime) == 2024) # Ambil data tahun 2024

if (nrow(data_2024) < 2) {
  cat("Error: Data 2024 tidak ditemukan atau tidak cukup untuk analisis.")
} else {
  cat(paste("Data 2024 berhasil difilter. Jumlah observasi:", nrow(data_2024), "\n\n"))

  # Ambil urutan keadaan (state) dari data 2024
  state_sequence_2024 <- data_2024$state

  cat("Contoh data 2024 yang digunakan:\n")
  kable(head(data_2024), caption = "Data Cuaca Tahun 2024 (6 Baris Pertama)")
}

## Data 2024 berhasil difilter. Jumlah observasi: 244
## Contoh data 2024 yang digunakan:
```

Data Cuaca Tahun 2024 (6 Baris Pertama)

Datetime total_rainfall_harianstate

```
2024-01-01 7.6Hujan Ringan
2024-01-02 17.4Hujan Ringan
2024-01-03 41.14Hujan Sangat Lebat
2024-01-04 3308.2Hujan Sangat Lebat
2024-01-05 1150.0Hujan Sangat Lebat
2024-01-06 12.8Hujan Ringan
```

```
cat("Ringkasan Statistik untuk Total Curah Hujan Harian (mm/hari) 2024:\n")
```

```
## Ringkasan Statistik untuk Total Curah Hujan Harian (mm/hari) 2024:
```

```
ringkasan_tutorial <- data_2024 %>%
  select(total_rainfall_harian) %>%
  summary()

kable(ringkasan_tutorial, caption = "Statistik Deskriptif Curah Hujan Harian 2024")
```

Statistik Deskriptif

Curah Hujan Harian

2024

total_rainfall_harian

Min.: 0.0

1st Qu.: 0.0

Median: 20.6

Mean: 345.6

3rd Qu.: 272.4

Max.: 5812.4

```
cat("Tabel Frekuensi Kategori Cuaca (State) 2024:\n")
```

```
## Tabel Frekuensi Kategori Cuaca (State) 2024:
```

```
# Hitung frekuensi dan persentase
urutan_state_logis <- c("Berawan", "Hujan Ringan", "Hujan Sedang", "Hujan Lebat", "Hujan Sangat Lebat") # Sesuai
an ini

tabel_frekuensi <- data_2024 %>%
  count(state, name = "Jumlah_Hari") %>%
  mutate(Persentase = round((Jumlah_Hari / sum(Jumlah_Hari)) * 100, 2)) %>%
  # Mengatur ulang urutan state sesuai urutan logis
  mutate(state = factor(state, levels = urutan_state_logis)) %>%
  arrange(state)

# Tampilkan tabel menggunakan kable
kable(tabel_frekuensi,
      caption = "Tabel Frekuensi Kategori Cuaca Harian 2024",
      col.names = c("Kategori (State)", "Jumlah Hari", "Persentase (%)"))
```

Tabel Frekuensi Kategori Cuaca Harian 2024

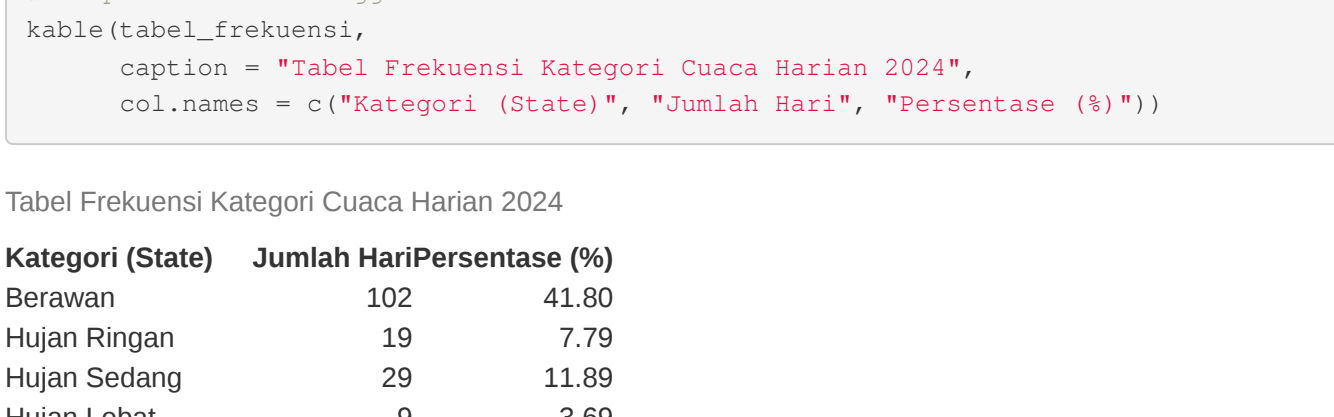
Kategori (State)	Jumlah Hari	Persentase (%)
Berawan	102	41.80
Hujan Ringan	19	7.79
Hujan Sedang	29	11.89
Hujan Lebat	9	3.69
Hujan Sangat Lebat	85	34.84

```
# Atur tema default untuk semua plot
theme_set(theme_minimal())

# Bar Chart untuk Frekuensi Kategori
plot_bar <- ggplot(tabel_frekuensi, aes(x = state, y = Jumlah_Hari, fill = state)) +
  geom_bar(stat = "identity", show.legend = FALSE) +
  geom_text(aes(label = paste0(Jumlah_Hari, " hari")), vjust = -0.5, size = 3) + # Tambah label jumlah
  labs(title = "Distribusi Kategori Cuaca (State) 2024",
        x = "Kategori (State)",
        y = "Jumlah Hari (Frekuensi)") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Putar label jika panjang

print(plot_bar)
```

Distribusi Kategori Cuaca (State) 2024

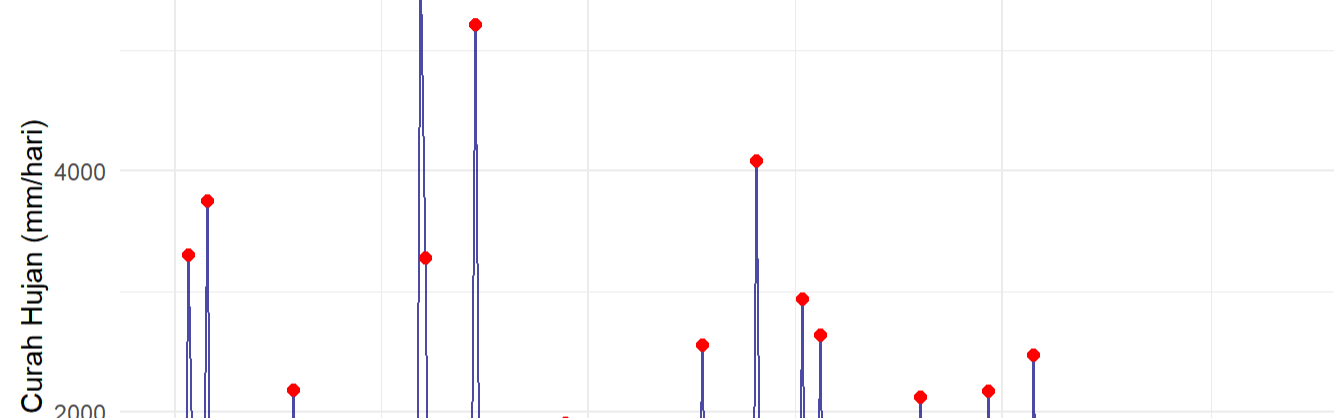


```
# Grafik Runtun Waktu (Time Series Plot)
plot_ts <- ggplot(data_2024, aes(x = DateTime, y = total_rainfall_harian)) +
  geom_line(color = "navy", alpha = 0.7) +
  # Tandai hari dengan hujan lebat (misal > 50mm)
  geom_point(data = filter(data_2024, total_rainfall_harian > 50),
             aes(x = DateTime, y = total_rainfall_harian),
             color = "red", size = 2) +
  labs(title = "Curah Hujan Harian Sepanjang Tahun 2024",
        subtitle = "Titik merah menandakan hujan lebat (> 50 mm/hari)",
        x = "Tanggal",
        y = "Total Curah Hujan (mm/hari)")

print(plot_ts)
```

Curah Hujan Harian Sepanjang Tahun 2024

Titik merah menandakan hujan lebat (> 50 mm/hari)



2. Identifikasi Ruang Keadaan (State Space)

Ruang keadaan adalah himpunan semua status cuaca unik yang terjadi pada tahun 2024.

```
unique_states <- unique(state_sequence_2024)
cat("Ruang Keadaan (State Space) 2024:\n")

## Ruang Keadaan (State Space) 2024:

print(unique_states)

## [1] "Hujan Ringan"      "Hujan Sangat Lebat" "Hujan Sedang"
## [4] "Hujan Lebat"      "Berawan"
```

3. Matriks Transisi

Kita mengestimasi matriks transisi 1-langkah (P) dari data. P_{ij} adalah probabilitas berpindah dari keadaan i ke keadaan j pada hari berikutnya.

```
# Mengestimasi matriks transisi dari data 2024
mc_fit_2024 <- markovchainFit(data = state_sequence_2024)

# Matriks probabilitas transisi P(1)
transition_matrix_2024 <- mc_fit_2024$estimate@transitionMatrix

cat("Matriks Transisi (P) 2024:\n")

## Matriks Transisi (P) 2024:
```

```
kable(round(transition_matrix_2024, 3), caption = "Matriks Transisi Probabilitas (P) 1-Langkah")
```

Matriks Transisi Probabilitas (P) 1-Langkah

	Berawan	Hujan Lebat	Hujan Ringan	Hujan Sangat Lebat	Hujan Sedang
Berawan	0.673	0.030	0.030	0.208	0.059
Hujan Lebat	0.333	0.111	0.111	0.000	0.444
Hujan Ringan	0.474	0.000	0.053	0.368	0.105
Hujan Sangat Lebat	0.165	0.035	0.141	0.482	0.176
Hujan Sedang	0.276	0.069	0.034	0.552	0.069

```
# Membuat objek markovchain formal
mc_cuaca_2024 <- new("Markovchain",
                     states = unique_states,
                     transitionMatrix = transition_matrix_2024,
                     name = "Model Cuaca 2024")

cat("\nRingkasan Model Rantai Markov:\n")

##
## Ringkasan Model Rantai Markov:
```

```
summary(mc_cuaca_2024)
```

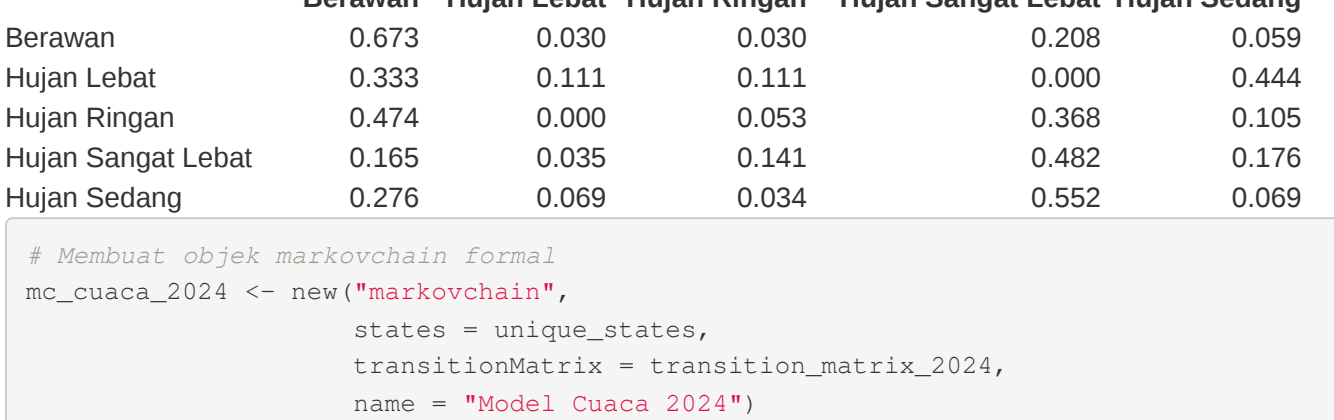
```
## Model Cuaca 2024 Markov chain that is composed by:
## Closed classes:
## Recurrent classes:
## (Hujan Ringan,Hujan Sangat Lebat,Hujan Sedang,Hujan Lebat,Berawan)
## Transient classes:
## NONE
## The Markov chain is irreducible
## The absorbing states are: NONE
```

4. Diagram Transisi

Visualisasi dari matriks transisi. Node adalah keadaan, dan panah adalah probabilitas transisi.

```
# Plot diagram transisi
tryCatch({
  plot(mc_cuaca_2024,
       main = "Diagram Transisi Status Cuaca (2024)",
       edge.arrow.size = 0.5,
       vertex.label.cex = 0.8)
}, error = function(e) {
  cat("Gagal membuat plot diagram transisi.")
})
```

Diagram Transisi Status Cuaca (2024)



5. Distribusi Stasioner

Ini adalah probabilitas jangka panjang (π) untuk berada di setiap keadaan, terlepas dari keadaan awal.

```
cat("Distribusi Stasioner (Steady States) 2024:\n")

## Distribusi Stasioner (Steady States) 2024:
```

```
tryCatch({
  stationary_dist <- steadyStates(mc_cuaca_2024)
  kable(t(round(stationary_dist, 4))), caption = "Distribusi Stasioner (Probabilitas Jangka Panjang)"
}, error = function(e) {
  cat("Tidak dapat menghitung distribusi stasioner (mungkin rantai tidak ergodik).\n")
})
```

Distribusi Stasioner

(Probabilitas Jangka Panjang)

Hujan Ringan	0.4213
Hujan Sangat Lebat	0.0372
Hujan Sedang	0.0739
Hujan Lebat	0.3487
Berawan	0.1191

6. Klasifikasi Ruang Keadaan

Menganalisis properti dari rantai ini (apakah terhubung, berulang, dll.).

```
cat("Analisis Klasifikasi Ruang Keadaan 2024:\n")
```

Analisis Klasifikasi Ruang Keadaan 2024:

```
# Cek Irreducibility
is_irreducible <- is_irreducible(mc_cuaca_2024)
cat(paste("Unik. Apakah Rantai Irreducible? (Semua state terhubung).", is_irreducible))

1. Apakah Rantai Irreducible? (Semua state terhubung): TRUE
```

```
# Cek Periode (1 = Aperiodic)
period_rantai <- period(mc_cuaca_2024)
cat(paste("n1. Periode Rantai (jika 1, berarti Aperiodic).", period_rantai))

2. Periode Rantai (jika 1, berarti Aperiodic): 1
```

```
if (is_irreducible && period_rantai == 1) {
  cat("\nnKesimpulan: Rantai ini adalah Ergodic (Irreducible & Aperiodic).")
  cat("\nnArtinya, Distribusi Stasioner valid dan unik.\n")
} else {
  cat("\nnKesimpulan: Rantai ini TIDAK ergodic. Analisis stasioner mungkin tidak berlaku.\n")
}
```

Kesimpulan: Rantai ini adalah Ergodic (Irreducible & Aperiodic). Artinya, Distribusi Stasioner valid dan unik.

```
# Cek state menyerap (Absorbing) atau sementara (Transient)
cat("\nn3. State Menyerap (Absorbing States):\n")

3. State Menyerap (Absorbing States):

print(absorbingStates(mc_cuaca_2024))

character(0)
```

```
cat("\nn4. State Sementara (Transient States):\n")
```

4. State Sementara (Transient States):

```
print(transientStates(mc_cuaca_2024))

character(0) ### Analisis:
```

• State Menyerap (Absorbing States): character(0)

Artinya: Model Anda tidak memiliki state menyerap. State menyerap adalah state yang jika sekali dimasuki, tidak akan pernah bisa ditinggalkan (probabilitas untuk tetap di state itu adalah 1). Dalam konteks cuaca, ini masuk akal. Tidak ada kondisi cuaca (seperti "Kering" atau "Hujan Lebat") yang jika terjadi, akan berlangsung selamanya tanpa pernah berubah.

• State Sementara (Transient States): character(0)

Artinya: Model Anda tidak memiliki state sementara. State sementara adalah state yang pada akhirnya akan Anda tinggalkan dan tidak akan pernah kembali lagi.

7. Simulasi

7.1 Probabilitas Langkah ke-n (P^n)

Ini adalah simulasi probabilitas berpindah dari keadaan i ke j dalam n langkah (hari). Mari kita hitung untuk $n = 3$.

```
# Menghitung probabilitas 3-langkah (P^3)
n_langkah <- 3
mc_3_langkah <- mc_cuaca_2024 ^ n_langkah

# Mengambil matriksnya
p_3_langkah <- mc_3_langkah@transitionMatrix

cat(paste("Matriks Transisi", n_langkah, "-Langkah (P^n):\n"))

## Matriks Transisi 3 -langkah (P^n):

kable(round(p_3_langkah, 3),
      caption = paste("Matriks Transisi Probabilitas", n_langkah, "-Langkah (P^n)"))
```

Matriks Transisi Probabilitas 3 -Langkah (P^n)

	Berawan	Hujan Lebat	Hujan Ringan	Hujan Sangat Lebat	Hujan Sedang
Berawan	0.467	0.036	0.066	0.321	0.110
Hujan Lebat	0.422	0.039	0.075	0.340	0.124
Hujan Ringan	0.426	0.036	0.074	0.346	0.118
Hujan Sangat Lebat	0.376	0.038	0.082	0.376	0.128
Hujan Sedang	0.389	0.038	0.078	0.371	0.125

```
# Mengambil probabilitas spesifik
tryCatch({
  prob_contoh <- p_3_langkah["Berawan", "Hujan Lebat"]
  cat(sprintf(
    "\nContoh: Probabilitas %d hari lagi 'Hujan Lebat', jika hari ini 'Berawan', adalah: %.4f (%.2f%%)\n",
    n_langkah, prob_contoh, prob_contoh * 100
  ))
}, error = function(e) {
  cat("\nnTidak dapat mengambil contoh probabilitas (pastikan nama state 'Berawan' dan 'Hujan Lebat' sesuai).\n")
})

##
## Contoh: Probabilitas 3 hari lagi 'Hujan Lebat', jika hari ini 'Berawan', adalah: 0.0363 (3.63%)
```

7.2 Simulasi Trajektori (Prediksi Urutan Cuaca)

Untuk menghasilkan simulasi urutan cuaca selama n hari ke depan, dengan asumsi kondisi awal tertentu.

```
# Simulasi 10 hari ke depan
n_days_simulasi <- 10
# Tentukan kondisi awal (hari ini)
initial_state <- "Hujan Sangat Lebat" #

simulasi_n_hari <- rmarkovchain(n = n_days_simulasi,
                                object = mc_cuaca_2024,
                                i0 = initial_state,
                                parallel = TRUE)

cat(paste("Hasil Simulasi", n_days_simulasi, "Hari ke Depan (dimulai dari '", initial_state, "'):\n"))

## Hasil Simulasi 10 Hari ke Depan (dimulai dari ' Hujan Sangat Lebat '):

print(simulasi_n_hari)
```

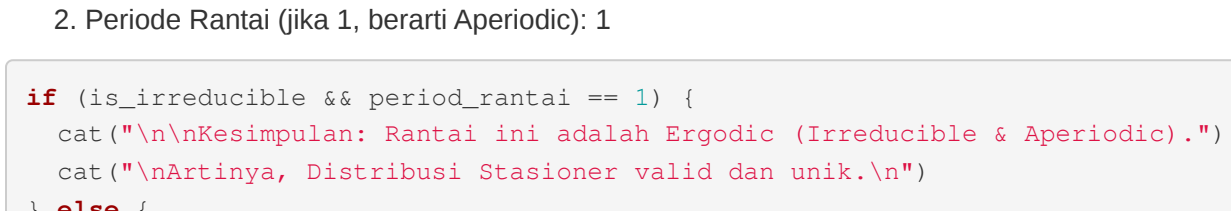
```
## [1] "Berawan"      "Hujan Ringan" "Hujan Ringan" "Hujan Ringan" "Hujan Ringan"
## [6] "Hujan Ringan" "Hujan Ringan" "Hujan Ringan" "Hujan Lebat"  "Hujan Ringan"
```

```
par(mfrow = c(1:1, 4:1, 4:1, 2:1))

# 2. Buat plot dengan 'las = 2' (vertikal)
plot(factor(simulasi_n_hari, levels = unique_states),
     main = paste("Simulasi", n_days_simulasi, "Hari Cuaca (Start: '", initial_state, "')"),

     # Gunakan label yang sudah diperbaiki,
     ylab = "Jumlah Hari (Frekuensi)",
     col = "lightblue",
     ylim = c(0, n_days_simulasi),
     las = 2)
```

Simulasi 10 Hari Cuaca (Start: ' Hujan Sangat Lebat ')



```
par(mfrow = c(5:1, 4:1, 4:1, 2:1))
```